

# On actions on cubic stochastic matrices

Irene Paniello <sup>1</sup>

*Department of Statistics and Operative Research*

*Public University of Navarre*

*31006 Pamplona, Spain*

E-mail: irene.paniello@unavarra.es

**Keywords:** Stochastic matrix, Stochastic group, Group action, Marginal distribution, Bivariate Markov chain.

**2010 Mathematics Subject Classification:** 15B51, 20M30, 60J10.

## Abstract

We consider the set of  $(n \times n \times n)$  cubic stochastic matrices of type (1,2) together with different multiplication rules that not only retain their stochastic properties but also endow this set with an associative semigroup structure. Then we introduce different actions of the semigroup of nonnegative column stochastic  $n \times n$  matrices on the set of cubic stochastic matrices of type (1,2) and study how these actions translate to the cubic matrix slices and marginal distributions. Actions introduced here provide an algebraic framework where considering different changes affecting the transition probabilities ruling certain biological populations.

## 1 Introduction.

Many discrete (physical, biological,...) dynamical systems can be modeled using stochastic processes. Then interactions occurring within the system between its particles (also the system evolution between its states) can be assumed to be ruled by a set of transition probabilities, oftentimes arranged into a square  $n \times n$  matrix. However sometimes, depending on the complexity of the system interactions (mainly when interactions involving more than

---

<sup>1</sup>Partially supported by the Spanish Ministerio de Ciencia y Tecnología and FEDER (MTM2013-45588-C3-2-P)

two particles may occur), multi-index matrices, also called tensors, need to be introduced. Independently of their dimension, all these matrices gathering transition probabilities will be nonnegative and endowed with some stochastic property.

Introducing multi-index stochastic matrices turns out to be necessary, for instance, when algebraically studying the backward inheritance (from progeny to ancestors) in Mendelian genetic populations, as introduced by Tian and Li [30]. These authors considered the problem of mathematically formalizing the dynamics of the transference of the genetic information through generations in a biological population ruled by Mendel's laws, focusing on the direction from progeny (offspring) to ancestors. This question was later revisited in [21], taking advantage of Maksimov's work [20] and rewritten in terms of  $(n \times n \times n)$  cubic matrices. Assuming that the probabilities ruling the genetic inheritance remain constant through generations and only a finite number of different genetic types exists, we can identify such a population's behavior with that of a time homogeneous discrete dynamical system.

In the late 40's Geiringer [11] remarked that the treatment of the problems arising in biological structures should be comparable to that of any other dynamical system, whenever the underlying stochastic processes behave similarly. This paper is aimed at providing an algebraic treatment of those stochastic tensors (multidimensional arrays) appearing in the study of backwards genetic inheritance in Mendelian genetic populations [21, 30]. More precisely, we focus on studying cubic stochastic matrices and their connection to square stochastic matrices.

Cubic stochastic matrices of type (1,2) were first considered by Maksimov in [20] in an attempt to explain processes in kinetic theory not representable in terms of Markov chains. Maksimov mainly focused on time homogeneous systems, but also considered Markov interaction processes (M.i.p.) [20, page 62] involving nonnegative stochastic square  $(n \times n)$  matrices together to cubic  $(n \times n \times n)$  stochastic matrices of type (2,3) acting simultaneously.

A similar connection to that occurring in M.i.p. between square and cubic matrices was previously considered by Lyubich [18, page 57] in terms of matrices of mutation rates acting on genetic systems. Lyubich studied the convergence of the so-called evolutionary operators, defined as the composition between a selection and a mutation operator.

The algebraic approach to products between stochastic tensors of different dimensions we attempt here is also motivated by Braman [3], who examined actions of (non-necessarily stochastic)  $n \times n \times n$  cubic matrices on  $n \times n$  square matrices. In the current work, however, it becomes necessary to take into account the stochastic nature of the matrices involved, a fact that clearly delimits which multiplications can be defined since we just look for those retaining stochastic properties.

After this introductory section, we first review different notions of square and cubic matrices, all having stochastic significance, with the aim of setting some notation. In the third section, we recall different types of cubic stochastic matrices, as originally introduced by Maksimov in [20], to focus on the so-called cubic stochastic matrices of type (1,2). These cubic stochastic matrices can gather up more complex transition probabilities than the usual square stochastic matrices, but also require of multiplication rules able to retain their stochastic properties. Therefore we generalize Maksimov's defined multiplications for stochastic matrices in an attempt to obtain new tools for studying cubic matrices. The new multiplication rules come from skewing, or more precisely weighting, the original Maksimov's definition.

In the fourth section we define different actions of the semigroup  $NS(n, \mathbb{R})$ , of nonnegative column stochastic  $n \times n$  matrices on the set  $CS_{(1,2)}(n, \mathbb{R})$  of cubic stochastic matrices of type (1,2), endowed with one of the previously defined weighted multiplications. We review how these actions translate to different lower-dimensional matrices related to cubic stochastic matrices of type (1,2), such as their slices and marginal distributions, as well as how they act on the bivariate Markov chains defined by the marginal distributions.

Finally, in the last section, besides ending up with a few remarks on some extensions that can be derived from these types of actions, we briefly describe the problem posed by Tian and Li in [20] from which our interest in actions defined on cubic matrices stems.

## 2 Square stochastic matrices and Markov chains.

This section is focused on the properties of the set of (nonnegative) column stochastic matrices, stressing on their connection to Markov processes. Examples over different fields are given, before restricting our attention to matrices with real entries, aimed by their role as transition probabilities ruling stochastic processes.

### 2.1 Square stochastic matrices.

A (square) *stochastic matrix* is a  $n \times n$  matrix  $\mathbf{P} = (p_{ij})_{i,j=1}^n$  such that each column sum equals to one

$$\sum_{i=1}^n p_{ij} = 1, \quad \text{for all } j = 1, \dots, n.$$

This definition corresponds to the so-called column stochastic matrices. Row stochastic matrices can be defined similarly. Vectors and matrices will be written in boldface capital letters ( $\mathbf{A}, \mathbf{v}, \dots$ ) and their usual multiplication simply by juxtaposition.

**2.2 Lemma.** *The following results hold for stochastic matrices over any arbitrary field  $F$ :*

- (i) *The product of two stochastic matrices is stochastic.*
- (ii) *The set of stochastic matrices forms a multiplicative semigroup.*
- (iii) *The set of nonsingular stochastic matrices forms a (multiplicative) group, called the stochastic group and denoted by  $S(n, F)$ .*

*Proof.* (i) is straightforward and (ii) and (iii) follow from considering the set of stochastic matrices as a subset of  $M(n, F)$ , the set of all  $n \times n$  matrices with entries in  $F$ .  $\square$

### 2.3 Example.

The set  $S(2, \mathbb{R})$  of  $2 \times 2$  nonsingular real stochastic matrices, i.e.,

$$S(2, \mathbb{R}) = \left\{ \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix} \mid a, b \in \mathbb{R}, a+b \neq 1 \right\},$$

is a group, whose Lie geometry is described in [27].

### 2.4 Examples.

In [26] examples of stochastic groups are given for different fields  $F$ :

- (i) Given an arbitrary field  $F$ ,  $S(n, F) \cong \text{Aff}(n-1, F)$  for all  $n \geq 2$ , where  $\text{Aff}(n-1, F)$  denotes the affine group, i.e. the set of all mappings  $f : F^{n-1} \rightarrow F^{n-1}$  such that  $f(\mathbf{x}^T) = \mathbf{A}\mathbf{x}^T + \mathbf{b}^T$ , with  $\mathbf{A} \in GL(n-1, F)$ , the general linear group, and  $\mathbf{x}, \mathbf{b} \in F^{n-1}$ . Then  $S(n, F)$  is a group with the composition.
- (ii) Given a Galois field  $F$  of order  $q = p^m$ ,  $p$  prime,  $S(n, F)$  includes examples such as  $S(2, 2) \cong \mathbb{Z}_2$  or  $S(2, 4) = A_4$  among many others. Here, following [26], we denote  $S(n, F)$  by  $S(n, q)$ .

**2.5 Remark.** The Markov Lie group was defined by Johnson in [13] to be the subset of the general linear group  $GL(n, \mathbb{R})$  preserving the linear form  $\sum_{i=1}^n x_i$ . As noted in [13, p. 254] the Markov Lie group is isomorphic to the affine group  $\text{Aff}(n-1, \mathbb{R})$ .

More general notions of stochastic matrices can be defined and also endowed with an algebraic structure.

## 2.6 Example.

Let  $F$  be  $\mathbb{R}$  or  $\mathbb{C}$ . An  $r$ -generalized doubly stochastic matrix is a matrix having all its row and column sums equal to  $r \in F$ . The set of all  $n \times n$   $r$ -generalized stochastic matrices is denoted by  $\Omega^r(n, F)$  being  $\Omega(n, F) = \cup_r \Omega^r(n, F)$  the set of all generalized doubly stochastic matrices. Nonnegative matrices in  $\Omega^1(n, F)$ , i.e. with  $r = 1$ , are called *doubly stochastic matrices*. Moreover (see [19]):

- (i)  $\Omega^r(n, F)$  is a subalgebra of  $M(n, \mathbb{C})$  iff  $r = 0$ .
- (ii)  $GL_1(n, F) = GL(n, F) \cap \Omega^1(n, F)$ , where  $GL(n, F)$  denotes the general linear group over  $F$ , is a Lie group with Lie algebra  $\Omega^0(n, F)$ , simple of dimension  $(n - 1)^2$  over  $F$ .

Doubly stochastic matrices were already considered by Sagle in [29] to prove that the set of doubly stochastic matrices over a field  $F$  of characteristic zero is a semisimple associative algebra generated by permutation matrices. Moreover it is isomorphic to  $F \oplus M(n - 1, F)$ .

Interesting subsets of stochastic matrices appear when matrix entries are considered in ordered fields.

## 2.7 Example.

Let  $NS(2, \mathbb{R})$  be the set of  $2 \times 2$  (real) stochastic matrices with nonnegative entries, i.e.,

$$NS(2, \mathbb{R}) = \left\{ \begin{pmatrix} 1-a & b \\ a & 1-b \end{pmatrix} \mid a, b \in \mathbb{R}, 0 \leq a, b \leq 1 \right\}.$$

Matrices in  $NS(2, \mathbb{R})$  are also called stochastic Markov matrices. These matrices provide transition probability matrices in models appearing when studying phylogenetic theory [27]. Sumner, Fernández-Sánchez and Jarvis considered more general Markov models in [28]. We recall here that Markov models are well-defined subsets of matrices of  $M(n, \mathbb{C})$  with all column sums equal to one.

Since this paper aimed to study those stochastic matrices arising when dealing with transition probabilities, from now on we will restrict to stochastic matrices with nonnegative (real) entries. Different results on the structure of the set of nonnegative matrices can be found in [4, 8, 25].

**2.8 Remark.** We will denote by  $NS(n, \mathbb{R})$  the set of all nonnegative real column stochastic  $n \times n$  matrices. Notice that we do not require matrices in  $NS(n, \mathbb{R})$  to be nonsingular.

The following results are straightforward.

**2.9 Lemma.** *Any convex linear combination of nonnegative stochastic matrices is stochastic, i.e., given  $\mathbf{A}, \mathbf{B} \in NS(n, \mathbb{R})$ , then  $\lambda \mathbf{A} + (1 - \lambda) \mathbf{B} \in NS(n, \mathbb{R})$  for any  $\lambda \in \mathbb{R}$  such that  $0 \leq \lambda \leq 1$ .*

**2.10 Proposition.**  *$NS(n, \mathbb{R})$  is a multiplicative semigroup.*

As it is broadly known, matrices in  $NS(n, \mathbb{R})$  are related to Markov processes.

## 2.11 Markov chains.

A *Markov chain* is a discrete time stochastic process  $X = \{X_t, t = 0, 1, \dots\}$  taking values on a finite or countably infinite state space  $\{1, 2, \dots\}$  satisfying the Markov property

$$P(X_{t+1} = i_{t+1} \mid X_0 = i_0, X_1 = i_1, \dots, X_t = i_t) = P(X_{t+1} = i_{t+1} \mid X_t = i_t),$$

where  $i_k \in \{1, 2, \dots\}$  for  $k = 0, 1, \dots, t+1$ . Homogeneous Markov processes are those whose transition probabilities  $P(X_{t+1} = i_{t+1} \mid X_t = i_t)$  do not depend on time  $t$ , i.e.  $P(X_{t+1} = i_{t+1} \mid X_t = i_t) = P(X_{t+1+k} = i_{t+1} \mid X_{t+k} = i_t)$  for all  $k$ . Assuming the set of states to be finite  $\{1, 2, \dots, n\}$  and denoting the transition probabilities by  $p_{ij} = P(X_{t+1} = i \mid X_t = j)$ , any homogeneous Markov chain gives rise to a matrix  $\mathbf{P} = (p_{ij})_{i,j=1}^n$  in  $NS(n, \mathbb{R})$ , i.e., a column stochastic  $n \times n$  matrix with nonnegative real entries.

Despite of the wide range of dynamical systems whose evolution can be modeled using Markov chains, multivariate Markov models have also proved to be useful to study dynamical systems, mainly whenever it becomes necessary to simultaneously deal with multiple data sequences. This happens, for instance when modeling DNA sequences or when studying backwards genetic inheritance in Mendelian genetic systems [6, 22, 30].

## 2.12 Multivariate Markov models.

Consider  $s$  categorical data sequences, each one with  $n$  possible states  $\{1, 2, \dots, n\}$ . Following [6] we denote by  $\mathbf{X}_t^{(j)}$  the state probability distribution of the  $j$ -th sequence at time  $t$ . Then the multivariate Markov model is

$$\mathbf{X}_{t+1}^{(j)} = \sum_{k=1}^s \lambda_{jk} \mathbf{P}^{(jk)} \mathbf{X}_t^{(k)}, \quad j = 1, \dots, s,$$

where  $\lambda_{jk} \geq 0$ ,  $1 \leq j, k \leq s$  and  $\sum_{k=1}^s \lambda_{jk} = 1$  for  $j = 1, 2, \dots, s$ . Matricially  $\mathbf{X}_{t+1} = \mathbf{Q} \mathbf{X}_t$ , i.e.

$$\mathbf{X}_{t+1} = \begin{pmatrix} \mathbf{X}_{t+1}^{(1)} \\ \mathbf{X}_{t+1}^{(2)} \\ \vdots \\ \mathbf{X}_{t+1}^{(s)} \end{pmatrix} = \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} & \dots & \lambda_{1s} \mathbf{P}^{(1s)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} & \dots & \lambda_{2s} \mathbf{P}^{(2s)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{s1} \mathbf{P}^{(s1)} & \lambda_{s2} \mathbf{P}^{(s2)} & \dots & \lambda_{ss} \mathbf{P}^{(ss)} \end{pmatrix} \begin{pmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \\ \vdots \\ \mathbf{X}_t^{(s)} \end{pmatrix}$$

Notice  $\mathbf{P}^{(jk)}$  is the transition probability matrix from the states in the  $k$ -th sequence to those in the  $j$ -th sequence, and therefore, an element in  $NS(n, \mathbb{R})$ . However column sums  $\mathbf{Q}$  need not be equal to one.

In the following sections we will consider the case  $s = 2$ , i.e., bivariate Markov chains as appear in [22] in connection to certain  $n \times n \times n$  matrices with stochastic properties.

## 2.13 Quadratic stochastic operators.

Prior to concluding this section it has to be pointed out that, regardless of their many applications, there are still many (biological, physical,...) systems that cannot be described by Markov models. See for instance



[2, 7, 12, 14, 15]. Another comprehensive reference is [10] where the stress is on the study of the dynamics of quadratic stochastic operators. As an example to illustrate one of these situations, let us consider quadratic stochastic operators acting on the  $(n - 1)$ -dimensional simplex in  $\mathbb{R}^n$ ,

$$S^{n-1} = \left\{ \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \geq 0, \sum_{i=1}^n x_i = 1 \right\}$$

i.e. transformations  $V : S^{n-1} \rightarrow S^{n-1}$  given by  $V(\mathbf{x})_k = \sum_{i,j=1}^n p_{ijk} x_i x_j$ ,  $k = 1, \dots, n$ , where

$$p_{ijk} \geq 0, \quad p_{ijk} = p_{jik} \quad \text{and} \quad \sum_{k=1}^n p_{ijk} = 1, \quad \forall i, j = 1, \dots, n.$$

Quadratic stochastic operators were introduced by Bernstein [5] in connection to mathematical models in genetics. Notice that matricially arranging the  $p_{ijk}$ 's we obtain not a square  $n \times n$  but a nonnegative cubic  $n \times n \times n$  matrix  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  whose entries have the following biological interpretation: In a genetic population with types  $\{1, 2, \dots, n\}$ ,  $p_{ijk}$  denotes the probability of a type  $k$  offspring to appear from parental types  $i$  and  $j$ . Here due to the assumption  $p_{ijk} = p_{jik}$ , parents's type order is not important.

### 3 Cubic stochastic matrices.

Multi-index arrays, usually called tensors, make it possible to matricially represent stochastic processes involving higher-order interaction systems. In this section we focus on third-order tensors, i.e., elements of  $\mathbb{R}^{n_1 \times n_2 \times n_3}$  to consider those having  $n_1 = n_2 = n_3$ . Following Maksimov's work [20], we will name these tensors to be cubic matrices (see also [16, 17]) and use boldface capital letters, e.g.  $\mathbf{P}$ , to denote them, as we did for vectors and matrices (equivalently tensors of order one and two respectively). Other different terms (rectangular matrices, cubical third-order tensors,..) may also found to refer to such matrices (see for instance [15]).

### 3.1 Cubic matrix.

A *cubic matrix*  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  is an object with three indices  $i, j, k$  which can be uniquely written in the form

$$\mathbf{P} = (p_{ijk})_{i,j,k=1}^n = \sum_{i,j,k=1}^n p_{ijk}(i, j, k),$$

where  $(i, j, k)$  denotes the unit cubic matrices, i.e.  $(i, j, k)$  is a  $n \times n \times n$  cubic matrix whose  $(i, j, k)$ th entry is equal to 1 and all other entries are equal to 0.

### 3.2 Cubic matrix decomposition.

Cubic matrices, as any other higher-dimensional tensor, give rise to different subarrays by fixing subsets of indices. Following the notation used in [15], we define the (horizontal, lateral and frontal) slices of a cubic matrix  $\mathbf{P}$  by fixing two indices and using the colon to indicate all elements (see [15, Figure 2.2]):

- (i) The *horizontal slices* of  $\mathbf{P}$  are  $\mathbf{P}_{i::} = (p_{ijk})_{j,k=1}^n$ , for all  $i = 1, \dots, n$ .
- (ii) The *lateral slices* of  $\mathbf{P}$  are  $\mathbf{P}_{:j:} = (p_{ijk})_{i,k=1}^n$ , for all  $j = 1, \dots, n$ .
- (iii) The *frontal slices* of  $\mathbf{P}$  are  $\mathbf{P}_{::k} = (p_{ijk})_{i,j=1}^n$ , for all  $k = 1, \dots, n$ .

Fixing every index but one, we obtain the following fibers (see [15, Figure 2.1]):

- (i) The *jk-columns* of  $\mathbf{P}$  are  $\mathbf{P}_{:jk} = (p_{1jk}, \dots, p_{njk})^T$ , for all  $j, k = 1, \dots, n$ .
- (ii) The *ik-rows* of  $\mathbf{P}$  are  $\mathbf{P}_{i:k} = (p_{i1k}, \dots, p_{ink})^T$ , for all  $i, k = 1, \dots, n$ .
- (iii) The *ij-tubes* of  $\mathbf{P}$  are  $\mathbf{P}_{ij:} = (p_{ij1}, \dots, p_{ijn})^T$ , for all  $i, j = 1, \dots, n$ .

Slices (and also fibers) can be used to unfold cubic matrices reordering the matrix entries into a rectangular matrix. This process is called *matrixization* or *unfolding* [15, 2.4]. To illustrate this procedure let  $\mathbf{P}$  be a  $3 \times 3 \times 3$  cubic matrix and let

$$\mathbf{P}_{::1} = \begin{pmatrix} p_{111} & p_{121} & p_{131} \\ p_{211} & p_{221} & p_{231} \\ p_{311} & p_{321} & p_{331} \end{pmatrix} \quad \mathbf{P}_{::2} = \begin{pmatrix} p_{112} & p_{122} & p_{132} \\ p_{212} & p_{222} & p_{232} \\ p_{312} & p_{322} & p_{332} \end{pmatrix}$$

and

$$\mathbf{P}_{::3} = \begin{pmatrix} p_{113} & p_{123} & p_{133} \\ p_{213} & p_{223} & p_{233} \\ p_{313} & p_{323} & p_{333} \end{pmatrix}$$

be the three frontal slices of  $\mathbf{P}$ . Then one of the possible matricizations of  $\mathbf{P}$  is the  $9 \times 3$  matrix given by its frontal slices:

$$\left( \mathbf{P}_{::1} \mid \mathbf{P}_{::2} \mid \mathbf{P}_{::3} \right) = \begin{pmatrix} p_{111} & p_{121} & p_{131} & p_{112} & p_{122} & p_{132} & p_{113} & p_{123} & p_{133} \\ p_{211} & p_{221} & p_{231} & p_{212} & p_{222} & p_{232} & p_{213} & p_{223} & p_{233} \\ p_{311} & p_{321} & p_{331} & p_{312} & p_{322} & p_{332} & p_{313} & p_{323} & p_{333} \end{pmatrix}$$

As noted in [15] the choice of slices or fibers when unfolding tensors is not as important as the reordering remaining consistent with respect to any defined computation.

### 3.3 Maksimov's cubic stochastic matrices.

A cubic matrix  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  is said to be:

- (i) *stochastic of type (1,2)* if  $p_{ijk} \geq 0$  and  $\sum_{i,j=1}^n p_{ijk} = 1$  for all  $k = 1, \dots, n$ .
- (ii) *stochastic of type (2,3)* if  $p_{ijk} \geq 0$  and  $\sum_{j,k=1}^n p_{ijk} = 1$  for all  $i = 1, \dots, n$ .
- (iii) *stochastic of type (1,3)* if  $p_{ijk} \geq 0$  and  $\sum_{i,k=1}^n p_{ijk} = 1$  for all  $j = 1, \dots, n$ .

These matrices were first introduced by Maksimov [20] to model physical systems such that interactions within the system could not be described in terms of Markov processes. Maksimov-like definitions can be extended to describe different classes of (always nonnegative) cubic stochastic matrices.

### 3.4 3-stochastic cubic matrices.

A cubic matrix  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  is said to be *3-stochastic* if  $p_{ijk} \geq 0$  and  $\sum_{k=1}^n p_{ijk} = 1$  for all  $i, j = 1, \dots, n$ . Clearly 3-stochastic cubic matrices are in a one-to-one correspondence to quadratic stochastic operators (see 2.13).

**3.5 Remark.** Even being  $\mathbf{P}$  cubic stochastic, neither its slices nor its fibers need to retain any stochastic property. Assume for instance that  $\mathbf{P}$  is cubic stochastic of type (1,2). Then only its frontal slices retain some stochastic nature. Indeed  $\{\mathbf{P}_{::k}\}_{k=1}^n$  is a family of bidimensional probability distributions on the set of pairs  $\{(i, j)\}_{i,j=1}^n$  (notice that  $\mathbf{P}_{::k}$  is nonnegative and  $\sum_{i,j=1}^n p_{ijk} = 1$ ,  $k = 1, \dots, n$ .)

**3.6 Remark.** The third order nature of cubic stochastic matrices increases the number of multiplications that can be defined for these matrices. See for instance [1] where the authors consider different associative multiplication rules defined on cubic matrices. Here we are only interested in multiplication rules that are consistent with the different definitions of stochasticity.

### 3.7 Maksimov · multiplication.

Maksimov defined in [20] the following multiplication for cubic stochastic matrices of type (1,2):

$$(i, j, k) \cdot (m, r, s) = \delta_{km}(i, j, s)$$

where  $(i, j, k)$  denotes the cubic matrix units and  $\delta_{km}$  is the Kronecker's delta. Elementwise, given  $\mathbf{A} = (a_{ijk})_{i,j,k=1}^n$ ,  $\mathbf{B} = (b_{ijk})_{i,j,k=1}^n$  and  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  such that  $\mathbf{P} = \mathbf{A} \cdot \mathbf{B}$ , we have

$$p_{ijs} = \sum_{k,r=1}^n a_{ijk} b_{krs} = \sum_{k=1}^n a_{ijk} b_{k+s}$$

where  $b_{k+s} = \sum_{r=1}^n b_{krs}$ .

**3.8 Proposition.** *The set of cubic stochastic matrices of type (1,2) with the · multiplication forms a convex semigroup (i.e. a multiplicative associative*

semigroup such that  $\lambda \mathbf{A} + (1 - \lambda) \mathbf{B}$  is again cubic stochastic of type (1,2) for any cubic stochastic matrices  $\mathbf{A}$  and  $\mathbf{B}$  of type (1,2) and any  $\lambda \in \mathbb{R}$  such that  $0 \leq \lambda \leq 1$ ).

*Proof.* The proof is analogous to that of [20, Proposition 3] taking into account that the associativity of  $\cdot$  is a straightforward checking.  $\square$

Some other multiplications were also introduced in [20] for different types of cubic stochastic matrices. Elementwise descriptions of these multiplication rules can be found in [16, 17]. The  $\cdot$  multiplication corresponds to one of the associative rules in [1].

**3.9 Remark.** To better understand how the  $\cdot$  multiplication acts on cubic stochastic matrices of type (1,2), let us recall the usual multiplication rule  $(i, j)(l, k) = \delta_{jl}(i, k)$  on nonnegative column square (i.e.  $n \times n$ ) matrices. Take  $\mathbf{P} = (p_{ij})_{i,j=1}^n \in NS(n, \mathbb{R})$  or consider, equivalently, the Markov chain with transition probability matrix  $\mathbf{P}$ . Then the  $(i, j)$ -th entry  $p_{ij}$  of  $\mathbf{P}$  gives the transition probability from a state  $j$  to a state  $i$  and the  $(i, j)$ -th entry  $p_{ij}^{(m)}$  of  $\mathbf{P}^{(m)}$ , the  $m$ -th power of  $\mathbf{P}$ , the probability that a state  $j$  reaches a state  $i$  after  $m$  transitions (i.e.  $m$ -step transition probabilities).

### 3.10 Probabilistic interpretation of the $\cdot$ multiplication.

A probabilistic interpretation of the  $\cdot$  multiplication appears in [20] in terms of transition probabilities for walking particles. Here, however, we will give a different interpretation in a biological setting. (See [9, 18, 21, 22].) Consider a biological system (or population) i.e. a collection of organisms of types  $\{1, 2, \dots, n\}$  closed with respect to reproduction, and let  $p_{ijk}$  be the probability for a type  $k$  individual to come from the mating of (ordered male and female) progenitors having types  $i$  and  $j$ . Then:

- (i)  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  is cubic stochastic of type (1,2) with

$$p_{ijk} = P(\text{father} = i, \text{mother} = j \mid \text{child} = k).$$

- (ii) The  $(i, j, k)$ -th entry of  $\mathbf{P}^{(m, \cdot)}$ ,  $m$ -th power of  $\mathbf{P}$  with respect to the  $\cdot$  multiplication, gives the probability for an individual of type  $k$  to

appear after  $m$  matings from an initial ordered pair of ancestors of types  $i$  and  $j$ .

More details of this approach will be given in the final section.

Maksimov reduced cubic matrices to lower dimensional square  $n \times n$  matrices obtained by adding the matrix entries along either the first or the second index.

### 3.11 Accompanying matrices. Marginal distributions.

Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be a cubic stochastic matrix of type (1,2). The following matrices were introduced in [20]:

- (a) The first accompanying matrix is  $\mathbf{P}_1 = (p_{i+k})_{i,k=1}^n$ , with  $p_{i+k} = \sum_{j=1}^n p_{ijk}$ .
- (b) The second accompanying matrix is  $\mathbf{P}_2 = (p_{+jk})_{j,k=1}^n$ , with  $p_{+jk} = \sum_{i=1}^n p_{ijk}$ .

These matrices were called  $i$ -accompanying and  $j$ -accompanying in [20] (and denoted  $\mathbf{P}_{(i)}$  and  $\mathbf{P}_{(j)}$  respectively), as they were later in [21] and [22]. However this notation, as first and second accompanying matrices, seems to better stress on how these matrices are formed. Clearly  $\sum_{i=1}^n p_{i+k} = \sum_{i=1}^n \sum_{j=1}^n p_{ijk} = \sum_{j=1}^n p_{+jk} = 1$ , thus both accompanying matrices are column stochastic. Hence, being also nonnegative, accompanying matrices are in  $NS(n, \mathbb{R})$ . We usually refer to both accompanying matrices of a cubic stochastic matrix of type (1,2) to be the *marginal distributions* of cubic matrix [22].

### 3.12 (1,2)-transposes.

Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be a cubic stochastic matrix of type (1,2). We define the (1,2)-transpose of  $\mathbf{P}$  and denote it by  $\mathbf{P}^{T(1,2)}$  to be the cubic matrix  $\mathbf{P}^{T(1,2)} = (q_{ijk})_{i,j,k=1}^n$  with  $q_{ijk} = p_{jik}$ , for all  $i, j, k = 1, \dots, n$ . We will say that  $\mathbf{P}$  is (1,2)-symmetric if  $\mathbf{P} = \mathbf{P}^{T(1,2)}$ , i.e.  $p_{ijk} = p_{jik}$  for all  $i, j, k = 1, \dots, n$ .

Although actions of (semi)groups on cubic stochastic matrices will be discussed later, we notice here that (1,2)-transpositions can be seen in terms of the action of  $S_3$ , the symmetric group of 3 elements, by permuting the matrix entries according to permutations on the index triples  $(i, j, k)$ . Then  $\mathbf{P}^{T(1,2)}$  results from the action of the transposition  $\sigma = (1, 2) \in S_3$  as follows:

$$\mathbf{P}^{T(1,2)} = \sigma \mathbf{P} = (p_{\sigma(i,j,k)})_{i,j,k=1}^n$$

where  $\sigma(i, j, k) = (j, i, k)$ .

**3.13 Proposition.** *Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be a cubic matrix. Then  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  is stochastic of type (1,2) if and only if so is  $\mathbf{P}^{T(1,2)}$ .*

*Proof.* It suffices to use the definition of cubic stochastic matrix of type (1,2) given in 3.3.  $\square$

**3.14 Lemma.** *Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be a cubic stochastic matrix of type (1,2). Then:*

- (i)  $(\mathbf{P}^{T(1,2)})_{h::} = \mathbf{P}_{:h:}$  for all  $h = 1, 2, \dots, n$ .
- (ii)  $(\mathbf{P}^{T(1,2)})_{:h:} = \mathbf{P}_{h::}$  for all  $h = 1, 2, \dots, n$ .
- (iii)  $(\mathbf{P}^{T(1,2)})_{::h} = (\mathbf{P}_{:h:})^T$  for all  $h = 1, 2, \dots, n$ .
- (iv) If  $\mathbf{P}$  is (1,2) symmetric, then its marginal distributions coincide.

*Proof.* See Lemma 1 and Lemma 2 in [23].  $\square$

Marginal distributions provide a different approach to the study cubic stochastic matrices of type (1,2). For instance, in [22] partial solutions to the ergodicity of these cubic matrices were obtained by considering the bivariate Markov model defined by the marginal distributions.

**3.15 Theorem.** [22, Theorem 1] *Let  $\mathbf{P}$  be a cubic stochastic matrix of type (1,2). Then*

$$\mathbf{X}_{t+1} = \begin{pmatrix} \mathbf{X}_{t+1}^{(1)} \\ \mathbf{X}_{t+1}^{(2)} \end{pmatrix} = \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} \end{pmatrix} \begin{pmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{pmatrix}$$

where  $\mathbf{P}^{(11)} = \mathbf{P}^{(12)}$  is the first accompanying matrix of  $\mathbf{P}$  and  $\mathbf{P}^{(21)} = \mathbf{P}^{(22)}$  the second accompanying matrix of  $\mathbf{P}$ , provides a bivariate Markov model whenever  $\lambda_{ij} \geq 0$  for  $1 \leq i, j \leq 2$  and  $\sum_{j=1}^2 \lambda_{ij} = 1$  for  $i = 1, 2$ .

Despite the fact that the relationship given in Theorem 3.15 between cubic stochastic matrices of type (1,2) and bivariate Markov matrices allows us to recover some information on the cubic matrices (see for instance results on cubic matrices ergodicity in [22, Theorem 2]), this connection is not enough to achieve a complete characterization of cubic matrices just looking at their marginal distributions. A similar lack of completeness follows when one considers the already mentioned cubic matrix multiplications, as neither the above introduced  $\cdot$  multiplication nor any of the remaining multiplication rules defined in [20], totally comprise the overall stochastic nature of the entries of a cubic stochastic matrix of type (1,2). To overcome drawbacks of the  $\cdot$  multiplication,  $(i, j, k) \cdot (m, r, s) = \delta_{km}(i, j, s)$ , only considering non-vanishing products (interactions) when the first index of the second (right) factor is involved and paying no attention to its middle index, the following multiplication was defined in [22] by weighting interactions involving also middle indices.

### 3.16 $\star$ multiplication.

The  $\star$  product of two cubic stochastic matrices of type (1,2) is defined as follows:

$$(i, j, k) \star (m, r, s) = \frac{1}{2}(\delta_{km} + \delta_{kr})(i, j, s).$$

**3.17 Remark.** Notice that

$$\begin{aligned} (i, j, k) \star (m, r, s) &= (i, j, k) \cdot \left( \frac{1}{2}(m, r, s) + \frac{1}{2}(m, r, s)^{T(1,2)} \right) = \\ &= (i, j, k) \cdot \left( \frac{1}{2}(m, r, s) + \frac{1}{2}(r, m, s) \right) = \\ &= \frac{1}{2}(\delta_{km} + \delta_{kr})(i, j, s). \end{aligned}$$

Thus it comes out that given cubic stochastic matrices  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  and  $\mathbf{Q} = (q_{ijk})_{i,j,k=1}^n$  of type (1,2), the  $(i, j, k)$ -th entry of  $\mathbf{P} \star \mathbf{Q}$  can be written



as:

$$\begin{aligned}
(\mathbf{P} \star \mathbf{Q})_{ijk} &= \frac{1}{2} \left( \sum_{r,s=1}^n p_{ijr} q_{rsk} + \sum_{r,s=1}^n p_{ijr} q_{srk} \right) = \frac{1}{2} \sum_{r=1}^n (p_{ijr} q_{r+k} + \sum_{s=1}^n p_{ijr} q_{+rk}) \\
&= \frac{1}{2} (p_{ij1}, \dots, p_{ijn}) \left( \begin{pmatrix} q_{1+k} \\ q_{2+k} \\ \vdots \\ q_{n+k} \end{pmatrix} + \begin{pmatrix} q_{+1k} \\ q_{+2k} \\ \vdots \\ q_{+nk} \end{pmatrix} \right) \\
&= (\mathbf{P}_{ij:})^T \left( \frac{1}{2} (\mathbf{Q}_1)_{:k} + \frac{1}{2} (\mathbf{Q}_2)_{:k} \right).
\end{aligned}$$

Hence  $(\mathbf{P} \star \mathbf{Q})_{ijk}$  is a equally weighted product of the  $ij$ -tube  $\mathbf{P}_{ij:}$  of  $\mathbf{P}$  and the  $k$ -th columns of the accompanying matrices of  $\mathbf{Q}$ . (Here we extend the notation introduced in 3.2 to square matrices.)

**3.18 Proposition.** *Let  $\mathbf{P}$  and  $\mathbf{Q}$  be cubic stochastic matrices of type  $(1,2)$ . Then  $\mathbf{P} \star \mathbf{Q}$  is again cubic stochastic of type  $(1,2)$ .*

*Proof.* Clearly  $\mathbf{P} \star \mathbf{Q}$  is a nonnegative matrix, hence it suffices to prove that  $\sum_{i,j=1}^n (\mathbf{P} \star \mathbf{Q})_{ijk} = 1$  for all  $k = 1, \dots, n$ , but being both  $\mathbf{P}$  and  $\mathbf{Q}$  cubic stochastic of type  $(1,2)$  we have

$$\begin{aligned}
\sum_{i,j=1}^n (\mathbf{P} \star \mathbf{Q})_{ijk} &= \sum_{i,j=1}^n \frac{1}{2} \sum_{r,s=1}^n (p_{ijr} q_{rsk} + p_{ijr} q_{srk}) \\
&= \frac{1}{2} \sum_{r,s=1}^n \left( \left( \sum_{i,j=1}^n p_{ijr} \right) q_{rsk} + \left( \sum_{i,j=1}^n p_{ijr} \right) q_{srk} \right) \\
&= \frac{1}{2} \sum_{r,s=1}^n (q_{rsk} + q_{srk}) = 1,
\end{aligned}$$

which implies that  $\mathbf{P} \star \mathbf{Q}$  is again cubic stochastic of type  $(1,2)$ .  $\square$

**3.19 Lemma.** *The following statements hold:*

- (i) *The  $\star$  multiplication is associative. Moreover cubic stochastic matrices of type  $(1,2)$  form a semigroup under  $\star$ .*
- (ii) *Let  $\mathbf{P}$  and  $\mathbf{Q}$  be cubic stochastic matrices of type  $(1,2)$ . If  $\mathbf{Q}$  is  $(1,2)$ -symmetric, then  $\mathbf{P} \star \mathbf{Q} = \mathbf{P} \cdot \mathbf{Q}$ .*

(iii) Let  $\mathbf{P}$  be a cubic stochastic matrix of type  $(1,2)$ . If  $\mathbf{P}$  is  $(1,2)$ -symmetric, then the  $m$ -th powers  $\mathbf{P}^{(m,\cdot)}$  and  $\mathbf{P}^{(m,\star)}$  of  $\mathbf{P}$ , with respect to  $\cdot$  and  $\star$  respectively, coincide.

*Proof.* The proof of (i) is just a case-by-case checking similar to that of Proposition 3.8. Write now  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  and  $\mathbf{Q} = (q_{ijk})_{i,j,k=1}^n$ . Then if  $q_{ijk} = q_{jik}$ , for all  $i, j, k = 1, \dots, n$ , we have:

$$\begin{aligned} (\mathbf{P} \star \mathbf{Q})_{ijk} &= \frac{1}{2} \sum_{r,s=1}^n (p_{ijr}q_{rsk} + p_{ijr}q_{srk}) \\ &= \frac{1}{2} \sum_{r,s=1}^n (p_{ijr}q_{rsk} + p_{ijr}q_{srk}) = \sum_{r,s=1}^n p_{ijr}q_{rsk} = (\mathbf{P} \cdot \mathbf{Q})_{ijk} \end{aligned}$$

Hence the  $(1,2)$ -symmetry of  $\mathbf{Q}$  ensures that  $\mathbf{P} \star \mathbf{Q} = \mathbf{P} \cdot \mathbf{Q}$ .

Finally to prove (iii) it suffices to note that  $m$ -th powers of  $\mathbf{P}$  with respect to both multiplication rules are well-defined, as a result of the associativity, and apply then (ii).  $\square$

Clearly the  $\star$  multiplication can be defined for arbitrary  $n \times n \times n$  matrices. But being mainly concerned with stochastic matrices, we have already restricted the definition to these matrices.

### 3.20 $\text{CS}_{(1,2)}(n, \mathbb{R})$ .

We will denote by  $\text{CS}_{(1,2)}(n, \mathbb{R})$  the set of all cubic stochastic matrices of type  $(1,2)$  with real entries and by  $(\text{CS}_{(1,2)}(n, \mathbb{R}), \star)$  the semigroup formed with the  $\star$  multiplication. Notice it is a (right) monoid, as  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$ , with  $p_{ijk} = 1$  if  $i = j = k$  and  $p_{ijk} = 0$  otherwise (see [15, Figure 2.4]) is cubic stochastic of type  $(1,2)$  and acts as a right (but not a left) identity element for  $\star$ .

### 3.21 Weighted multiplications.

The  $\star$  multiplication is a particular case of a more general weighted product

$$(i, j, k) \star_{(\lambda_1, \lambda_2)} (m, r, s) = (i, j, k) \cdot \left( \lambda_1(m, r, s) + \lambda_2(m, r, s)^{T(1,2)} \right)$$

for segregation coefficients  $\lambda_1, \lambda_2 \geq 0$ , such that  $\lambda_1 + \lambda_2 = 1$ . Elementwise, given  $\mathbf{P}, \mathbf{Q} \in \text{CS}_{(1,2)}(n, \mathbb{R})$ ,

$$\begin{aligned}
(\mathbf{P} \star_{(\lambda_1, \lambda_2)} \mathbf{Q})_{ijk} &= \lambda_1 \sum_{r,s=1}^n p_{ijr} q_{rsk} + \lambda_2 \sum_{r,s=1}^n p_{ijr} q_{srk} = \\
&= \sum_{r,s=1}^n p_{ijr} (\lambda_1 q_{rsk} + \lambda_2 q_{srk}) = \sum_{r=1}^n p_{ijr} (\lambda_1 q_{r+k} + \lambda_2 q_{+rk}) \\
&= (p_{ij1}, \dots, p_{ijn}) \left( \lambda_1 \begin{pmatrix} q_{1+k} \\ q_{2+k} \\ \vdots \\ q_{n+k} \end{pmatrix} + \lambda_2 \begin{pmatrix} q_{+1k} \\ q_{+2k} \\ \vdots \\ q_{+nk} \end{pmatrix} \right) \\
&= (\mathbf{P}_{ij:})^T (\lambda_1 (\mathbf{Q}_1)_{:k} + \lambda_2 (\mathbf{Q}_2)_{:k}).
\end{aligned}$$

Clearly  $\star$  arises in the equally weighted case  $\lambda_1 = \lambda_2 = \frac{1}{2}$ , so that  $\star = \star_{(\frac{1}{2}, \frac{1}{2})}$ , whereas Maksimov's  $\cdot$  multiplication corresponds to the case  $(\lambda_1, \lambda_2) = (1, 0)$ , i.e.  $\cdot = \star_{(1,0)}$ .

With a similar proof to that of Lemma 3.19 the following results follow.

**3.22 Proposition.** *Let  $\lambda_1, \lambda_2$  be nonnegative real numbers such that  $\lambda_1 + \lambda_2 = 1$  and  $\mathbf{P}, \mathbf{Q} \in \text{CS}_{(1,2)}(n, \mathbb{R})$ . Then:*

- (i)  $\mathbf{P} \star_{(\lambda_1, \lambda_2)} \mathbf{Q}$  is again in  $\text{CS}_{(1,2)}(n, \mathbb{R})$ .
- (ii) The  $\star_{(\lambda_1, \lambda_2)}$  multiplication is associative and the set of cubic stochastic matrices of type  $(1,2)$  form a monoid under the multiplication  $\star_{(\lambda_1, \lambda_2)}$ .
- (ii) If  $\mathbf{Q}$  is  $(1,2)$ -symmetric, then  $\mathbf{P} \star_{(\lambda_1, \lambda_2)} \mathbf{Q} = \mathbf{P} \cdot \mathbf{Q}$ .
- (iii) If  $\mathbf{P}$  is  $(1,2)$ -symmetric, then  $\mathbf{P}^{(m, \cdot)} = \mathbf{P}^{(m, \star_{(\lambda_1, \lambda_2)})}$ , where  $\mathbf{P}^{(m, \cdot)}$  and  $\mathbf{P}^{(m, \star_{(\lambda_1, \lambda_2)})}$  denote the  $m$ -th powers of  $\mathbf{P}$  with respect to  $\cdot$  and  $\star_{(\lambda_1, \lambda_2)}$  respectively.

**3.23**  $(\text{CS}_{(1,2)}(n, \mathbb{R}), \star_{(\lambda_1, \lambda_2)})$ .

For any nonnegative  $\lambda_1, \lambda_2$  such that  $\lambda_1 + \lambda_2 = 1$ , we will denote by  $(\text{CS}_{(1,2)}(n, \mathbb{R}), \star_{(\lambda_1, \lambda_2)})$  the monoid formed by the set  $\text{CS}_{(1,2)}(n, \mathbb{R})$  with the

$\star_{(\lambda_1, \lambda_2)}$  multiplication. Recall that the case  $\lambda_1 = \lambda_2 = \frac{1}{2}$  corresponds to  $(\text{CS}_{(1,2)}(n, \mathbb{R}), \star)$ .

Despite the different weightings that can be considered on cubic matrices, our main interest will continue focused on the equally weighted  $\star$  case. This product is the one arisen from the problem originally motivating our interest in cubic matrices, due to their role when studying the transference of genetic inheritance in backwards Mendelian populations (see [21, 30]).

Products involving tensors having different dimensions were already considered in [3], where stress was on the action of  $n \times n \times n$  tensors on  $n \times n$  matrices. In the following section we will deal with an opposite problem, as we will consider actions on cubic stochastic matrices by square matrices.

## 4 Actions on cubic stochastic matrices.

In this section we consider different actions of the semigroup (of nonnegative column square stochastic matrices)  $NS(n, \mathbb{R})$  on the set  $\text{CS}_{(1,2)}(n, \mathbb{R})$  of cubic stochastic matrices of type (1,2), paying also attention to how these actions behave on slices and marginal distributions of the cubic matrices.

### 4.1 Group actions.

We recall that a group  $(G, \circ)$  acts on a set  $S$  if  $g \cdot s \in S$  is defined for all  $g \in G$  and  $s \in S$  and we have  $e \cdot s = s$  and  $(g_1 \circ g_2) \cdot s = g_1 \cdot (g_2 \cdot s) \in S$  for all  $s \in S$  and  $g_1, g_2 \in G$ , where  $e$  denotes the identity of  $G$ .

A similar definition holds if  $G$  is a semigroup with identity element. All semigroups appearing in this section have identity element. Notice that if  $G$  is just a semigroup, as not every element  $g \in G$  needs to have an inverse  $g^{-1}$  in  $G$ , not all actions by elements of  $G$  on  $S$  will be reversible, but just those given by invertible elements.

**4.2 Theorem.** *Let  $\{(i, j)\}_{i,j=1}^n$  and  $\{(i, j, k)\}_{i,j,k=1}^n$  be the square and cubic matrix units. Then, extended by linearity,*

$$(i) \ (i, j) \otimes_1 (r, s, t) = \delta_{jr}(i, s, t)$$

$$(ii) \ (i, j) \otimes_2 (r, s, t) = \delta_{js}(r, i, t)$$

define actions of the semigroup  $NS(n, \mathbb{R})$  on the set  $CS_{(1,2)}(n, \mathbb{R})$  of cubic stochastic matrices of type  $(1,2)$ .

**4.3 Remark.** We notice that  $\otimes_2$  can be obtained from  $\otimes_1$  (and vice versa) by conjugation given by the  $(1,2)$ -transposition. Indeed, denoting by  $T(1,2)$  the  $(1,2)$ -transposition on  $CS_{(1,2)}(n, \mathbb{R})$ , we have

$$(i, j) \otimes_2 (r, s, t) = \left[ (i, j) \otimes_1 [(r, s, t)^{T(1,2)}] \right]^{T(1,2)}$$

thus  $\otimes_2 = T(1,2) \circ \otimes_1 \circ T(1,2)$ , or equivalently  $T(1,2) \circ \otimes_2 = \otimes_1 \circ T(1,2)$ , considering here  $(1,2)$ -transposes as the result of the action of the transposition  $(1,2) \in S_3$  on the cubic matrix indices (see 3.12).

**4.4 Remark.** Actions  $\otimes_1$  and  $\otimes_2$  can be alternatively defined using direct product notation as follows:

$$(i) \ ((i, j) \times 1 \times 1) \otimes_1 (r, s, t) = \delta_{jr}(i, s, t)$$

$$(ii) \ (1 \times (i, j) \times 1) \otimes_2 (r, s, t) = \delta_{js}(r, i, t)$$

where tensor product components act on the different indices of the triple  $(r, s, t)$ . Thus, given  $\mathbf{A} = (a_{ij})_{i,j=1}^n \in NS(n, \mathbb{R})$  and  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n \in CS_{(1,2)}(n, \mathbb{R})$ , we have

$$(i) \ \mathbf{A} \otimes_1 \mathbf{P} = (\mathbf{A} \times \mathbf{I}_n \times \mathbf{I}_n) \cdot \mathbf{P} = \left( \sum_{r=1}^n a_{ir} p_{rst} \right)_{i,s,t=1}^n$$

$$(ii) \ \mathbf{A} \otimes_2 \mathbf{P} = (\mathbf{I}_n \times \mathbf{A} \times \mathbf{I}_n) \cdot \mathbf{P} = \left( \sum_{s=1}^n a_{is} p_{rst} \right)_{r,i,t=1}^n$$

where  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix. Using this alternative approach to  $\otimes_1$  and  $\otimes_2$ , as actions of  $NS(n, \mathbb{R}) \times NS(n, \mathbb{R}) \times NS(n, \mathbb{R})$  on the set of cubic matrices, Theorems 4.2 and 4.5 follow easily. However detailed proofs are also included for completeness. Moreover, see Remark 4.3, we have  $(\mathbf{A} \otimes_2 \mathbf{P})^{T(1,2)} = \mathbf{A} \otimes_1 (\mathbf{P}^{T(1,2)})$ , where, as noted before,  $T(1,2)$  can be understood as the action of the transposition  $(1,2) \in S_3$  on the cubic matrix indices.

*Proof.* Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be a cubic stochastic matrix of type (1,2) and let  $\mathbf{A} = (a_{ij})_{i,j=1}^n$  be in  $NS(n, \mathbb{R})$  (so we have  $\sum_{i=1}^n a_{ij} = 1$  for all  $j = 1, \dots, n$ ). We first claim that both  $\mathbf{A} \otimes_1 \mathbf{P}$  and  $\mathbf{A} \otimes_2 \mathbf{P}$  are cubic stochastic of type (1,2). Indeed, if we write  $\mathbf{A} \otimes_1 \mathbf{P} = (q_{ist})_{i,s,t=1}^n$ , then  $\mathbf{A} \otimes_1 \mathbf{P}$  has nonnegative entries  $q_{ist} = \sum_{r=1}^n a_{ir} p_{rst}$  satisfying

$$\sum_{i,s=1}^n q_{ist} = \sum_{i,s=1}^n \sum_{r=1}^n a_{ir} p_{rst} = \sum_{r,s=1}^n \left( \sum_{i=1}^n a_{ir} \right) p_{rst} = \sum_{r,s=1}^n p_{rst} = 1.$$

Hence  $\mathbf{A} \otimes_1 \mathbf{P}$  is cubic stochastic of type (1,2). Similarly we prove that  $\mathbf{A} \otimes_2 \mathbf{P} = (w_{rit})_{r,i,t=1}^n$  with  $w_{rit} = \sum_{s=1}^n a_{is} p_{rst}$  is cubic stochastic of type (1,2). (Proposition 3.13 and Remark 4.3 also apply here.)

Clearly  $\mathbf{I}_n \otimes_1 \mathbf{P} = \mathbf{P} = \mathbf{I}_n \otimes_2 \mathbf{P}$  and, therefore, it only remains to be checked that

$$(i, j) \otimes_1 \left[ (r, s) \otimes_1 (t, u, v) \right] = \left[ (i, j)(r, s) \right] \otimes_1 (t, u, v),$$

and

$$(i, j) \otimes_2 \left[ (r, s) \otimes_2 (t, u, v) \right] = \left[ (i, j)(r, s) \right] \otimes_2 (t, u, v),$$

hold. Hence both  $\otimes_1$  and  $\otimes_2$  define actions of  $NS(n, \mathbb{R})$  on  $CS_{(1,2)}(n, \mathbb{R})$ .  $\square$

Taking into account the semigroup structure of both  $NS(n, \mathbb{R})$  and  $CS_{(1,2)}(n, \mathbb{R})$  the next theorem follows.

**4.5 Theorem.**  $\otimes_1$  and  $\otimes_2$  define actions of the semigroup  $NS(n, \mathbb{R})$  of nonnegative column square stochastic matrices on  $(CS_{(1,2)}(n, \mathbb{R}), \star)$ .

The above defined actions act independently on  $(CS_{(1,2)}(n, \mathbb{R}), \star)$  and coincide on (1,2)-symmetric matrices.

**4.6 Proposition.** Actions  $\otimes_1$  and  $\otimes_2$  commute on the set of cubic stochastic matrices of type (1,2), i.e.:

$$(i, j) \otimes_1 \left( (k, l) \otimes_2 (m, r, s) \right) = (k, l) \otimes_2 \left( (i, j) \otimes_1 (m, r, s) \right).$$

Hence  $\mathbf{B} \otimes_2 (\mathbf{A} \otimes_1 \mathbf{P}) = \mathbf{A} \otimes_1 (\mathbf{B} \otimes_2 \mathbf{P})$  for any stochastic matrices  $\mathbf{A}, \mathbf{B} \in NS(n, \mathbb{R})$  and  $\mathbf{P} \in CS_{(1,2)}(n, \mathbb{R})$ .

*Proof.* Using Remark 4.4 it suffices to notice that  $(\mathbf{A} \times \mathbf{I}_n \times \mathbf{I}_n)(\mathbf{I}_n \times \mathbf{B} \times \mathbf{I}_n) = (\mathbf{I}_n \times \mathbf{B} \times \mathbf{I}_n)(\mathbf{A} \times \mathbf{I}_n \times \mathbf{I}_n)$  in  $NS(n, \mathbb{R}) \times NS(n, \mathbb{R}) \times NS(n, \mathbb{R})$ .  $\square$

**4.7 Corollary.** *Let  $\mathbf{P}$  be a cubic stochastic matrix of type  $(1,2)$ . If  $\mathbf{P}$  is  $(1,2)$ -symmetric, then  $\mathbf{A} \otimes_2 \mathbf{P} = (\mathbf{A} \otimes_1 \mathbf{P})^{T(1,2)}$  for any  $\mathbf{A} \in NS(n, \mathbb{R})$ .*

*Proof.* The result follows from Remark 4.3.  $\square$

Next we deal with the interaction between the  $\otimes_1$  and  $\otimes_2$  actions and the semigroup structure  $(CS_{(1,2)}(n, \mathbb{R}), \star)$  of the set of cubic stochastic matrices of type  $(1,2)$  under the  $\star$  multiplication.

**4.8 Proposition.** *Actions  $\otimes_1$  and  $\otimes_2$  are consistent with the  $\star$  multiplication of cubic stochastic matrices of type  $(1,2)$ :*

$$\begin{aligned} (i) \quad (i, j) \otimes_1 \left( (m, r, s) \star (p, q, k) \right) &= \left( (i, j) \otimes_1 (m, r, s) \right) \star (p, q, k) \\ (ii) \quad (i, j) \otimes_2 \left( (m, r, s) \star (p, q, k) \right) &= \left( (i, j) \otimes_2 (m, r, s) \right) \star (p, q, k) \end{aligned}$$

*Proof.* The proof is a straightforward checking.  $\square$

As noted in Remark 3.5, the family of frontal slices  $\{\mathbf{P}_{::k}\}_{k=1}^n$  of a cubic stochastic matrix  $\mathbf{P}$  of type  $(1,2)$  gives rise to a family of bidimensional probability distributions. (Notice this is in fact equivalent to  $\mathbf{P}$  being stochastic of type  $(1,2)$ .) Actions  $\otimes_1$  and  $\otimes_2$  act on the frontal slices of  $\mathbf{P}$  independently.

**4.9 Theorem.** *Let  $\mathbf{P}$  be a cubic stochastic matrix of type  $(1,2)$  and  $\mathbf{A} \in NS(n, \mathbb{R})$ . Then the following statements hold:*

$$\begin{aligned} (i) \quad (\mathbf{A} \otimes_1 \mathbf{P})_{::k} &= \mathbf{A} \mathbf{P}_{::k}. \\ (ii) \quad (\mathbf{A} \otimes_2 \mathbf{P})_{::k} &= \mathbf{A} (\mathbf{P}_{::k})^T = \mathbf{A} (\mathbf{P}^{T(1,2)})_{::k}. \end{aligned}$$

*Proof.* Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  and  $\mathbf{A} = (a_{ij})_{i,j=1}^n$ . The above equalities follow from noticing that  $\mathbf{A} \otimes_1 \mathbf{P} = (q_{ijk})_{i,s,k=1}^n$  with  $q_{isk} = \sum_{r=1}^n a_{ir} p_{rsk}$  and  $\mathbf{A} \otimes_2 \mathbf{P} = (w_{rik})_{r,i,k=1}^n$  with  $w_{rik} = \sum_{s=1}^n a_{is} p_{rsk}$ . The additional equality in (ii) follows from Lemma 3.14(iii).  $\square$

**4.10 Corollary.** *The matricization of  $\mathbf{A} \otimes_1 \mathbf{P}$  by its frontal slices in*

$$\left( \mathbf{A} \mathbf{P}_{::1} \mid \dots \mid \mathbf{A} \mathbf{P}_{::n} \right)$$

*and that of  $\mathbf{A} \otimes_2 \mathbf{P}$  is*

$$\left( \mathbf{A}(\mathbf{P}_{::1})^T \mid \dots \mid \mathbf{A}(\mathbf{P}_{::n})^T \right)$$

*Proof.* It suffices to use Theorem 4.9 and recall the matricization of cubic matrices given in 3.2.  $\square$

The role marginal distributions have to better understand cubic stochastic matrices lead us to study how actions  $\otimes_1$  and  $\otimes_2$  behave on their marginal distributions and also of the derived bivariate Markov chain (see Theorem 3.15).

**4.11 Proposition.** *Let  $\mathbf{P}$  be a cubic stochastic matrix of type (1,2).*

(i) *The  $\otimes_1$  action is consistent first accompanying matrices, whereas  $\otimes_2$  is consistent with second accompanying matrices, i.e., for any  $\mathbf{A} \in NS(n, \mathbb{R})$  we have:*

$$(a) \quad (\mathbf{A} \otimes_1 \mathbf{P})_1 = \mathbf{A} \mathbf{P}_1,$$

$$(b) \quad (\mathbf{A} \otimes_2 \mathbf{P})_2 = \mathbf{A} \mathbf{P}_2.$$

(ii) *Second accompanying matrices remain invariant under  $\otimes_1$ , whereas first accompanying matrices remain invariant under  $\otimes_2$ , i.e.:*

$$(a) \quad (\mathbf{A} \otimes_1 \mathbf{P})_2 = \mathbf{P}_2,$$

$$(b) \quad (\mathbf{A} \otimes_2 \mathbf{P})_1 = \mathbf{P}_1.$$

*Proof.* Let  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  be cubic stochastic of type (1,2) and  $\mathbf{A} = (a_{ij})_{i,j=1}^n \in NS(n, \mathbb{R})$ . Write  $\mathbf{A} \otimes_1 \mathbf{P} = (q_{ist})_{i,s,t=1}^n$ . To prove (i) we notice that  $(\mathbf{A} \otimes_1 \mathbf{P})_1 = \mathbf{A} \mathbf{P}_1$  follows from

$$q_{i+t} = \sum_{s=1}^n q_{ist} = \sum_{s=1}^n \sum_{r=1}^n a_{ir} p_{rst} = \sum_{r=1}^n a_{ir} \left( \sum_{s=1}^n p_{rst} \right) = \sum_{r=1}^n a_{ir} p_{r+t}$$

and similarly one proves that  $(\mathbf{A} \otimes_2 \mathbf{P})_2 = \mathbf{A} \mathbf{P}_2$  holds.



Next to prove  $(\mathbf{A} \otimes_1 \mathbf{P})_2 = \mathbf{P}_2$  it suffices to notice that

$$q_{+st} = \sum_{i=1}^n q_{ist} = \sum_{i=1}^n \sum_{r=1}^n a_{ir} p_{rst} = \sum_{r=1}^n p_{rst} \left( \sum_{i=1}^n a_{ir} \right) = \sum_{r=1}^n p_{rst} = p_{+st},$$

since  $\sum_{i=1}^n a_{ir} = 1$ , for all  $r = 1, \dots, n$ . Therefore the second accompanying matrix of  $\mathbf{A} \otimes_1 \mathbf{P}$  is exactly that of  $\mathbf{P}$  and  $(\mathbf{A} \otimes_2 \mathbf{P})_1 = \mathbf{P}_1$  follows analogously.  $\square$

**4.12 Theorem.** Let  $\mathbf{P}$  be a cubic stochastic matrix of type  $(1,2)$  with associated bivariate Markov chain  $\mathbf{X}_{t+1} = \mathbf{Q}\mathbf{X}_t$  where

$$\mathbf{Q} = \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} \end{pmatrix}$$

being  $\mathbf{P}^{(11)} = \mathbf{P}^{(12)}$  the first accompanying matrix  $\mathbf{P}_1$  of  $\mathbf{P}$  and  $\mathbf{P}^{(21)} = \mathbf{P}^{(22)}$  the second accompanying matrix  $\mathbf{P}_2$  of  $\mathbf{P}$  and  $\lambda_{ij} \geq 0$  for  $1 \leq i, j \leq 2$  and  $\sum_{j=1}^2 \lambda_{ij} = 1$  for  $i = 1, 2$ . For any nonnegative stochastic matrix  $\mathbf{A} \in NS(n, \mathbb{R})$ :

(i) The  $\otimes_1$  action of  $\mathbf{A}$  on  $\mathbf{P}$  gives rise to the bivariate Markov chain  $\mathbf{X}_{t+1} = \mathbf{Q}_1 \mathbf{X}_t$  with

$$\mathbf{Q}_1 = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} \end{pmatrix}$$

(ii) The  $\otimes_2$  action of  $\mathbf{A}$  on  $\mathbf{P}$  gives rise to the bivariate Markov chain  $\mathbf{X}_{t+1} = \mathbf{Q}_2 \mathbf{X}_t$  with

$$\mathbf{Q}_2 = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} \end{pmatrix}$$

(iii) The simultaneous  $\otimes_1$  and  $\otimes_2$  actions of  $\mathbf{A}$  on  $\mathbf{P}$  give rise to the bivariate Markov chain  $\mathbf{X}_{t+1} = \mathbf{Q}_3 \mathbf{X}_t$  with

$$\mathbf{Q}_3 = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \lambda_{11} \mathbf{P}^{(11)} & \lambda_{12} \mathbf{P}^{(12)} \\ \lambda_{21} \mathbf{P}^{(21)} & \lambda_{22} \mathbf{P}^{(22)} \end{pmatrix}$$

*Proof.* (i) follows from Proposition 4.11 as  $\mathbf{A}\mathbf{P}^{(11)} = \mathbf{A}\mathbf{P}^{(12)} = \mathbf{A}\mathbf{P}_1 = (\mathbf{A} \otimes_1 \mathbf{P})_1$  whereas  $(\mathbf{A} \otimes_1 \mathbf{P})_2 = \mathbf{P}_2$ . Analogously one proves (ii) and (iii).  $\square$

Results obtained in this section, besides on their own interest to study the structure of the semigroup of cubic stochastic matrices of type (1,2) through actions by  $NS(n, \mathbb{R})$ , will be applied in the following section in the particular setting of those populations arising in backwards Mendelian genetic inheritance. The algebraic formulation of this problem is given in [30] (see also 3.10).

## 5 An example and further remarks.

In this section, prior to gathering some concluding remarks, we recover the problem posed by Tian and Li [30] on the backwards inheritance of the genetic information in populations ruled by Mendel's laws. Actions on cubic stochastic matrices are used to show how inheritance rules are modified when different mutations (i.e. changes modifying the inheritance probabilities) occur.

### 5.1 Modeling mutations in genetic populations.

Let  $\mathbf{P}$  be a cubic stochastic matrix of type (1,2). In [21] these cubic matrices were revisited in connection to coalgebras with genetic realization (or genetic coalgebras for short). A *coalgebra with genetic realization* is a real coalgebra  $(C, \Delta)$ , i.e. a finite dimensional real vector space  $C$  with a basis  $\mathcal{B} = \{e_1, \dots, e_n\}$  and a linear map  $\Delta : C \rightarrow C \otimes C$ , called *comultiplication*, such that  $\Delta(e_k) = \sum_{i,j=1}^n \beta_{ij}^k e_i \otimes e_j$  with

- (i)  $0 \leq \beta_{ij}^k \leq 1$ , for all  $i, j, k = 1, \dots, n$ ,
- (ii)  $\sum_{i,j=1}^n \beta_{ij}^k = 1$  for all  $k = 1, \dots, n$ .

Genetic coalgebras were introduced in [30] to model the backwards (from progeny to ancestors) genetic inheritance in Mendelian genetic systems.

Looking at the elements of (the natural basis)  $\mathcal{B}$  as a complete set of representatives of the different types  $\{1, 2, \dots, n\}$  for a given hereditary trait, we can identify  $\beta_{ij}^k = P(\text{father} = i, \text{mother} = j \mid \text{child} = k)$ . Then

$\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$  with  $p_{ijk} = \beta_{ij}^k$ , for all  $i, j, k = 1, \dots, n$ , is cubic stochastic of type (1,2) [21].

Examples of genetic coalgebras, including a description of their related cubic stochastic matrices can be found in [23]. The reader is also referred to [24] for examples of graphical representations of these systems by oriented graphs.

One of the questions that remained unsolved in [30] was how to introduce mutations in Mendelian populations. Mutations have to be understood as changes affecting to the transition probabilities, (i.e. to the  $\beta_{ij}^k$ 's), so that the set  $\{\beta_{ij}^k = p_{ijk}\}_{i,j,k=1}^n$  is replaced by a new set of transition probabilities  $\{\tilde{p}_{ijk}\}_{i,j,k=1}^n$  still defining a cubic stochastic matrix. A first approach to this problem was given in [30, Theorem 5.7] by composing the coalgebra multiplication  $\Delta$  with different linear maps having the property of leaving the simplex  $S^{n-1}$  (i.e. the set of probability distributions on the set  $\{1, 2, \dots, n\}$ ) invariant.

We notice that, due to the correspondence between genetic coalgebras and the elements of  $\text{CS}_{(1,2)}(n, \mathbb{R})$  [21], the backwards evolution of the corresponding genetic population can be studied using cubic matrices. The  $\star$  multiplication defined in 3.16 assumes paternal and maternal heritage to be symmetrical (i.e. equally weighted). Then changes in the inheritance probabilities (i.e. mutations) are introduced by having nonnegative column stochastic square  $(n \times n)$  matrices, i.e. elements in  $NS(n, \mathbb{R})$ , acting on either the paternal  $(\otimes_1)$  or maternal  $(\otimes_2)$  heritage.

Take a fixed hereditary trait  $k$  and consider the corresponding frontal  $k$ -slice  $\mathbf{P}_{::k}$  of  $\mathbf{P}$ . Then  $\mathbf{P}_{::k}$  encloses the probability distribution that traces the genetic inheritance from a trait  $k$  back to its progenitors. Now it suffices to recall that having  $\mathbf{A} \in NS(n, \mathbb{R})$  acting on  $\mathbf{P}$ , amounts to have  $\mathbf{A}$  acting on all  $\mathbf{P}_{::k}$ ,  $k = 1, \dots, n$ , simultaneously (see Proposition 4.9).

On the other hand, taking into account that the first accompanying matrix  $\mathbf{P}_1 = (p_{i+k})_{i,k=1}^n$  encloses the probabilities  $p_{i+k} = P(\text{father} = i \mid \text{child} = k)$ , and that actions of  $\mathbf{A} \in NS(n, \mathbb{R})$  on  $\mathbf{P}$  result on actions (by multiplication) of  $\mathbf{A}$  on the marginal distributions on  $\mathbf{P}$  (see Proposition 4.11 and Theorem 4.12), it becomes possible to consider mutations affecting only one of the two (paternal or maternal) inheritance lines. Similar results can be achieved when considering the second accompanying matrix  $\mathbf{P}_2$  of  $\mathbf{P}$  since then  $p_{+jk} = P(\text{mother} = j \mid \text{child} = k)$ .

As a result the equally weighted  $\star$  multiplication, together to the actions  $\otimes_1$  and  $\otimes_2$ , defined on  $CS_{(1,2)}(n, \mathbb{R})$  provide a framework to consider different mutations in coalgebraic structures appearing in genetics.

## 5.2 Further remarks.

In this paper we deal with nonnegative stochastic  $(n \times n)$  square and  $(n \times n \times n)$  cubic matrices. We first put aside their stochasticity to consider different algebraic properties of these sets of matrices. It was then possible, with suitable defined multiplications, to obtain different semigroup structures. We have also defined different actions of the semigroup  $NS(n, \mathbb{R})$  of nonnegative column stochastic  $n \times n$  matrices on the semigroup  $(CS_{(1,2)}(n, \mathbb{R}), \star)$  of cubic stochastic matrices of type (1,2). For these actions, aimed by the role matricization (i.e. tensor decomposition into matrices [15]) and marginal distributions have to study these cubic matrices [23, 24], we have also considered how actions translate to their slices and related bivariate Markov chains.

Finally we have brought back the fact that our interest in cubic matrices stemmed from the search for an algebraic framework where mutations affecting to the transference of the information in backwards genetic inheritance (i.e. from progeny to ancestors) in Mendelian populations could be algebraically described. Keeping this in mind, we began this section briefly considering how mutations changing the inheritance probabilities can be identified to actions on cubic stochastic matrices of type (1,2) by elements of  $NS(n, \mathbb{R})$ .

We remark that, although we have just focused on those cubic matrices linked to Mendelian inheritance, similar approaches would also make sense for different types of stochastic tensors. It would suffice to consider an acting (semi)group consistent with, both dimension and stochastic properties of, the involved tensors. Consider, for instance, a 3-stochastic cubic matrix  $\mathbf{P} = (p_{ijk})_{i,j,k=1}^n$ . Then for any permutation  $\sigma \in S_n$ , i.e. any element of the symmetric group of  $n$  elements,  $\sigma\mathbf{P} = (p_{ij\sigma(k)})_{i,j,k=1}^n$  is again 3-stochastic. It would therefore become possible to define an action of  $S_n$  on the set of 3-stochastic cubic matrices by permuting the matrix frontal slices according to  $\sigma$ , or equivalently, actions on quadratic stochastic operators (see 2.13).

This type of actions is expected to have reasonable applications to study changes in transition probabilities, not only in biological populations but also in many other types of systems or processes, as they provide an algebraic framework where displaying how these changes modify the future evolution of the system under consideration.

## Acknowledgements.

The author wants to thank the referee for his/her carefully reading of the original manuscript and also for all provided suggestions that have undoubtedly contributed to improve the paper. The author also wants to thank Professor M. Ladra for his comments on a preliminary version of the manuscript.

## References

- [1] R. Bai, H. Liu, M. Zhang, 3-Lie algebras realized by cubic matrices, *Chin. Ann. Math. Ser. B* 35 (2), 261-270 (2014).
- [2] L. Boltzam, *Selected works*, Nauka, Moscow, 1984.
- [3] K. Braman, Third-order tensors as linear operators on a space of matrices, *Linear Algebra and its Applications* 433, 1241-1253 (2010).

- [4] D. R. Brown, On clans of non-negative matrices, *Proc. Amer. Math. Soc.* 15, 671-674 (1964).
- [5] S.N. Bernstein, Solution of a mathematical problem connected with the theory of heredity, *Ann. Sci. de l'Ukraine* 1, 83-114 (1924).
- [6] W. Ching, M.K. Ng, Markov chains: Models, Algorithms and Applications, International Series in Operations Research and Management Science, 83. Springer, New York (2006).
- [7] A. Dohtani, Occurrence of chaos in higher-dimensional discrete-time systems, *SIAM J. Appl. Math.* 52 (6), 1707-1721 (1992).
- [8] P. Flor, On groups of non-negative matrices, *Compositio Math.* 21, 376-382 (1969).
- [9] N. Ganikhodjaev, F. Mukhamedov, Ergodic properties of discrete quadratic stochastic processes defined on von Neumann algebras, *Izv. Math.* 64 (5), 873-890 (2000).
- [10] N. Ganikhodjaev, F. Mukhamedov, Quantum quadratic operators and processes. *Lecture Notes in Mathematics*, 2133. Springer 2015.
- [11] H. Geiringer, On some mathematical problems arising in the development of the Mendelian Genetics, *J. Amer. Stat. Assoc.* 44, 526-547 (1949).
- [12] R. D. Jenks, Quadratic differential systems for interactive population models, *J. Differential Equations* 5, 497-514 (1969).
- [13] J. E. Johnson, Markov-type Lie groups in  $GL(n, \mathbb{R})$ , *J. Math. Phys.* 26, 252-257 (1985).
- [14] Y. Kawamura, Structure of Cubic Matrix Mechanics, *Progr. Theoret. Phys.* 109, 1-10 (2003).
- [15] T. G. Kolda, B. W. Bader, Tensor decomposition and applications, *SIAM Review* 51 (3), 455-500 (2009).
- [16] M. Ladra, U. A. Rozikov, Flow of finite-dimensional algebras, *J. Algebra* 470, 263-288 (2017).
- [17] M. Ladra, U. A. Rozikov, Algebras of cubic matrices, *arXiv* 1606.02681.

- [18] Y. I. Lyubich, Mathematical structures in population genetics. Springer-Verlag, Berlin (1992).
- [19] B. Mourad, On a Lie-theoretic Approach to Generalized Doubly Stochastic Matrices and Applications, *Linear and Multilinear Algebra* 52 (2), 99-113 (2004).
- [20] V. M. Maksimov, Cubic stochastic matrices and their probability interpretation, *Theory Probab. Appl.* 41 (1), 55-69 (1997).
- [21] I. Paniello, Stochastic matrices arising from genetic inheritance, *Linear Algebra and its Applications* 434, 791-800 (2011).
- [22] I. Paniello, Marginal distributions on genetic coalgebras, *J. Math. Biology* 68 (5), 1071-1087 (2014).
- [23] I. Paniello, On evolution operators of genetic coalgebras, *J. Math. Biology* 74 (1-2), 149-168 (2017).
- [24] I. Paniello, Backwards genetic inheritance through coalgebra-graphs, *Linear and Multilinear Algebra* 65 (5), 943-961 (2017).
- [25] R. J. Plemmons, Regular nonnegative matrices, *Proc. Amer. Math. Soc.* 39 (1), 26-32 (1973).
- [26] D. G. Poole, The Stochastic Group, *The American Mathematical Monthly* 798-801 (1995).
- [27] J. G. Sumner, Lie geometry of  $2 \times 2$  Markov matrices, Letter to Editor, *Journal of Theoretical Biology* 327, 88-90 (2013).
- [28] J. G. Sumner, J. Fernández-Sánchez, P. D. Jarvis, Lie Markov models, *Journal of Theoretical Biology* 298, 16-31 (2012).
- [29] A. A. Sagle, Stochastic groups and algebras. Current trends in matrix theory (Auburn, Ala., 1986), 331-337, North-Holland, New York (1987).
- [30] J. Tian, B-L. Li, Coalgebraic structure of genetic inheritance, *Mathematical Biosciences and Engineering* 1 (2), 243-266 (2004).